

ABSTRACT OF THE DISCLOSURE

A technique is disclosed for disambiguating speech input for multimodal systems by using a combination of speech and visual I/O interfaces. When the user's speech input is not recognized with sufficiently high confidence, a the user is presented with a set of possible matches using a visual display and/or speech output. The user then selects the intended input from the list of matches via one or more available input mechanisms (e.g., stylus, buttons, keyboard, mouse, or speech input). These techniques involve the combined use of speech and visual interfaces to correctly identify user's speech input. The techniques disclosed herein may be utilized in computer devices such as PDAs, cellphones, desktop and laptop computers, tablet PCs, etc.